DOCUMENT RESUME

ED 212 991 CS 006 503

AUTHOR	Tierney, Robert J.; Mosenthal, James H.
TITLE	The Cohesion Concept's Relationship to the Coherence of Text. Technical Report No. 221.
INSTITUTION	Bolt, Beranek and Newman, Inc., Cambridge, Mass.; Illinois Univ., Urbana. Center for the Study of Reading.
SPONS AGENCY	National Inst. of Education (ED), Washington, D.C.
PUB DATE	Oct 81
CONTRACT	400-76-0116
NOTE	44p.
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	Analysis of Variance; *Coherence; *Cohesion (Written Composition); *Correlation; Factor Analysis; Grade 12; High Schools; *Reading Research; Secondary School Students; *Writing Research
IDENTIFIERS	*Text Structure; *Textual Analysis

ABSTRACT

Using M. A. K. Halliday's and R. Hasan's concept of cohesion, a study examined statistical accounting of cohesive ties as a means of measuring and evaluating text coherence. Twelve grade twelve students were provided outlines on two topics and asked to write essays for each topic. One group of six students was familiar with the topics, the other group of six students unfamiliar. Cohesive patterning was determined, based on a cohesive analysis of each text. A MANOVA revealed the effect of topic on cohesive patterning but showed no effect of familiarity on cohesion. When the essays were ranked according to their level of coherence and compared to the ordering of texts according to the cohesive analysis, no relation appeared between coherence ranking and cohesive patterning. Additional analyses of types of cohesive ties showed no significant relation between specific types of cohesive patterning and coherence. It was concluded that the cohesion of a text, as defined by Halliday and Hasan, bears no direct, causal relationship to the coherence of text. (RL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED212991

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)
This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.
Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

CENTER FOR THE STUDY OF READING

Technical Report No. 221

THE COHESION CONCEPT'S RELATIONSHIP TO THE COHERENCE OF TEXT

Robert J. Tierney and James H. Mosenthal
University of Illinois at Urbana-Champaign

October 1981

University of Illinois
at Urbana-Champaign
51 Gerty Drive
Champaign, Illinois 61820

Bolt Beranek and Newman Inc.
50 Moulton Street
Cambridge, Massachusetts 02238

The research reported herein was supported in part by the National Institute of Education under Contract No. HEW-NIE-C-400-76-0116.

CS206503

EDITORIAL BOARD

Paul Jose and Jim Mosenthal
Co-Editors

Harry Blanchard
Nancy Bryant
Larry Colker
Avon Crismore
Roberta Ferrara
Anne Hay
Asghar Iran-Nejad
Jill LaZansky
Ann Myers
Kathy Starr
Cindy Steinberg
William Tirre

Paul Wilson

Michael Nivens, Editorial Assistant

The Cohesion Concept's Relationship to the Coherence of Text

The purpose of the present study is to explore the extent to which Halliday and Hasan's cohesion concept (1976), applied as a text analysis system, serves as an index of textual coherence. Specifically, we wanted to check to what extent a statistical accounting of cohesive ties was a legitimate means of measuring and evaluating text coherence.

Our concern for the cohesion concept derives from our interest in structural analyses of text (see Tierney & Mosenthal, 1980). Text analysis systems are used to help predict and explain comprehension of text. The cohesion concept of Halliday and Hasan (1976) offered a special appeal since it claimed to represent a non-structural property of text. A non-structural, cohesive analysis of text seemed to offer a complementary, original means of examining the effect of text features on comprehension.

However, a general problem with the use of text analysis, and especially so with cohesion, is the assumption that the features of text subject to analysis cause or determine a text's coherence for a reader. The mistake, we feel, is to regard coherence as the product of textual features. Morgan (1978), Morgan and Sellner (1980), and Levy (1979) make this point with respect to the cohesion concept of Halliday and Hasan (1976). Their argument, quite simply, is that cohesion, used as a text analysis system, amounts to a counting and categorizing of words and phrases in text defined by Halliday and Hasan to be cohesive. Therefore, any remark about a text's

coherence based on a cohesion index is open to the criticism that coherence is being located in the text and described as a product of specifically textual features. The present study examines the legitimacy of this criticism and attempts to come to some conclusions about the relationship of cohesion and cohesive analysis to the coherence of text.

Method

Overview of the Design

Our design afforded comparisons of a range of writing samples written about two topics with content and structure held relatively constant for each topic. We chose the following two topics for our study--a biographical sketch of Nathaniel Hawthorne and a brief discussion of the theme of evil in Hawthorne's work. These topics were chosen because of their relevance to work done by the classes participating in the study. These topics were taken from the transcript of a cassette recording accompanying a filmstrip on the life and work of Hawthorne (Great Authors: Nathaniel Hawthorne, Schloat Productions, 1973). We made content and structure constant for the two topics by constructing outlines from the content of the transcript (students were asked to write essays on the biography and theme of evil topics using the outlines that we constructed).

In an effort to obtain a wide range of variability across student-generated essays, familiarity conditions were set up on the assumption that they would likely create such differences. The students were assigned to a familiar or unfamiliar group based on whether they were shown the

Hawthorne filmstrip. The familiar group saw the fimstrip and heard the cassette recording of the transcript which accompanied the filmstrip and from which the outlines were constructed. The unfamiliar group saw a filmstrip, accompanied by a cassette recording, on the topic of the political philosophy of Henry David Thoreau.

Our design represents a 2 x 2 repeated measures design. Our dependent measures consisted of proportions of types of cohesive ties used by the students in the essays that they wrote. Following the cohesive analyses, three teachers of college rhetoric courses rated the essays, within topics, with respect to clarity of expression and general coherence. These rankings were compared to the cohesive analyses.

Subjects

Twelfth grade students from two advanced rhetoric classes participated in the study. One class was arbitrarily designated to be the familiar group, while the other class was designated the unfamiliar group. Twelve out of 20 students, 6 for each class, completed all the work described in the Familiarization and Essay Production sections. The subject pool was limited to these 12 students.

The classes were taught by the same teacher who covered the same course content in each class. Scholastic Aptitude Verbal test scores for each student were obtained and the entire group of students was ranked according to these scores. Based on the distribution of students from both classes over this ranking it was concluded that the classes were of roughly equal ability and, therefore, no allowance was made for ability differences between

the familiar and unfamiliar groups (mean of familiar group = 47; mean of unfamiliar group = 45).

Materials

The procedure for constructing the outlines was based on a principle used by Meyer (1975) in the construction of her content diagrams--that is, the principle of using indentation to represent subordination. A TOPIC and DETAILS division were set up to emphasize subordinate relationships. Indentation only has subordinative significance within the TOPIC or DETAILS division within which it occurs. The students were given practice in reading such outlines using texts and outlines from a pilot study. Appendix A contains the outlines that the students used to write their essays.

Familiarization and Essay Production

Familiarization and essay production demanded four full 50-minute class periods. On day one the two filmstrips were shown, the Nathaniel Hawthorne filmstrip to the class chosen to be the familiar group, and the Thoreau filmstrip to the class chosen to be the unfamiliar group. Also, the two classes practiced interpreting the outline format using a text and an outline on the topic of the Gold Rush. On days two, three, and four, students were given outlines from which they were asked to generate four essays, two on Hawthorne and two on Thoreau. On these days, students were given the outlines one at a time (no student had two outlines at any one time), from individual student packets. To control for an ordering effect, the packets were assembled with no two outlines based on the same author

allowed to be adjacent. These packets were randomly assigned to the students. Thus, with respect to the Hawthorne tasks, the Thoreau tasks served as a means of partially randomizing the order of essay tasks presented to each student over the three days of writing. The students were directed to write an essay from each outline, imagining that their peers were their audience.

The Cohesion Analyses

A cohesion analysis was used based upon the system described by Halliday and Hasan in Cohesion in English (1976). The opening chapter of Cohesion in English provides a summary description of the technical aspects of the cohesion concept, as well as a theoretical justification for it. In their system a cohesive tie is defined by two cohesive elements, one presupposing and the other presupposed, crossing at least one sentence boundary. Typical ties consist of a pronoun, the presupposing item, and its referent, the presupposed item. In our cohesive analyses, the most important item of information identified the cohesive relationship between the presupposing and presupposed items. The relationship was categorized as referential, substitutive, elliptical, conjunctive, or lexical. Appendix B represents a simplified cohesive analysis of the first two paragraphs of one of the essays written on the topic of Hawthorne's biography.

After each text was analyzed, a table of the percentage figures for each cohesion category and subcategory was made, Substitution and ellipsis were left out of the analysis simply because there were insignficant numbers of their cohesive types used in any of the essays. The low

frequency of these cohesive types is consistent with Halliday and Hasan's contention that these two types of cohesive ties typify informal, conversational texts--whether written or spoken--more than they do expository text. Because they represented proportions, the percentage figures were transformed according to the angular or inverse sine transformation (see Kirk, 1968). Analyses were carried out on the transformed data to assess whether there were systematic differences across text topic and/or familiarity conditions with respect to the cohesive variables.

Coherence Rankings

The coherence rankings made by the rhetoric instructors were used to determine the relationship between essay coherence and cohesive patterning. Independent of the major statistical analyses, proportions of certain tie-types used were examined to see if they correlated with the coherence rankings. Each group of essays corresponding to one of the text topics was ranked by the three raters. The raters were asked simply to rate the essays in terms of their general level of coherence. All essays had been typed, with any spelling errors corrected, but with all syntax and punctuation left intact.

The procedure for determining an essay's overall rank within its group, based upon the three raters' responses, was as follows: If the raters agreed upon the essay's rank within one place, it was concluded that they agreed upon the essay's ranking. For example, if an essay was ranked 1 by the first rater, 2 by the second rater, and 3 by the third rater, it was concluded that there was 100% agreement on the general coherence level of the essay relative to the other essays in the same text group. If an essay

was ranked 1 by the first rater, 3 by the second rater, and 6 by the third rater then it was concluded that there was 67% agreement in the ranking of the essay (raters 1 and 2 agreeing). If an essay was ranked 1 by the first rater, 3 by the second rater and 5 by the third rater, it was concluded that there was 67% agreement for either of 2 rankings--a rank of from 1-3 or a rank of from 3-5. If no other essay competed for the rank of 3 then this was the rank assigned. If an essay did compete for the rank of 3 then the percentage of agreement among raters or the range of rankings were used to determine the final ranks of the two competing essays. If there was no agreement between the three raters, no ranking was given to an essay.

Results and Discussion

The Hawthorne essays were first examined to assess whether or not there was a topic or familiarity main effect for variation in cohesive patterning. This examination was carried out using a MANOVA followed up with a discriminant analysis. Aspects of the original proportion data were then compared with the coherence rankings to determine if there existed any relationship between relative coherence and cohesive patterning. The non-transformed proportion data are presented in Table 1.

Results of Variation in Cohesive Patterning Across Topic and Familiarity Conditions

The MANOVA and discriminant analysis, as well as univariate analyses (ANOVA), addressed the issue of variation across text topic and familiarity

conditions. Group means and standard deviations using the transformed scores are presented in Table 2. The MANOVA revealed a non-significant interaction effect for Familiarity x Topic, $F(3,8) = .652$, $p = .256$. However, there was a significant main effect for Topic, $F(3,8) = 25.166$, $p < .0002$. In other words, between the two text conditions studied, cohesion, as measured by proportions of reference, conjunction, and lexical ties used, significantly varied with respect to text topic. Since the trends were not uniform across topic, a discriminant analysis was conducted. The standardized discriminant function coefficients, for the one statistically significant discriminant function which was obtained, are as follows: 4.175-Reference Ties, 2.022-Conjunction Ties, 3.529-Lexical ties. These coefficients suggest that the reference and lexical cohesion categories were most powerful in discriminating cohesive patterning between topics.

Discussion of Variation in Cohesive Patterning Across Topic and Familiarity Conditions

The results of the MANOVA suggest a negative answer to the original question: To what extent is a statistical accounting of cohesive ties a legitimate means of measuring and evaluating text coherence? The MANOVA and discriminant analysis establish the effect of topic on cohesion as cohesion is defined by the integrated patterning of reference, conjunction, and lexical tie proportions per essay. In other words, topic accounts for most of the variation in cohesive categories. The significance of such an effect should not be exaggerated. It cannot be interpreted that given

a topic, a text's coherence is predicted by a correct proportion of reference, conjunction, and lexical ties. All that might be concluded, based on the MANOVA results, is that topic appears to affect the options a writer has for using cohesive items.

In their article, Discourse and Linguistic Theory (1980), Morgan and Sellner make a strong argument for the subordinate relationship of cohesive patterning to topic and to coherence--criticizing any interpretation that would attribute to cohesive patterning a priority in the determination of text coherence. They make the following statement:

> One might have assumed that the coherence of a text was a matter of content, which would have, of course, linguistic consequences. In a coherent biography of Churchill, for example, one would expect frequent mention of Churchill; one would therefore expect frequent occurrence of words like Churchill, he, him, his, and so on. The source of coherence would be in the content, and the repeated occurrence of certain words would be the consequence of content coherence, not something that was a source of coherence. It would be a serious mistake to construe this linguistic manifestation as cause, rather than effect (p. 25).

In other words, proportions of ties fall out of a coherent rendering of a topic. It is intuitively clear that a coherent rendering of a topic will not fall out of particular proportions of reference, conjunction, and lexical ties. One can't write an equation for coherence using proportions as weights for cohesive category variables.

Morgan and Sellner's argument is directly relevant to the MANOVA results reported for this study. The topic for essay 1 is the life of Hawthorne--and one does expect frequent use of words to refer to Hawthorne (such as Hawthorne, he, his, etc.). This helps explain the greater proportion of reference ties for the biography essay. The topic for essay 2 is the theme of evil in the work of Hawthorne--here one expects the use of words to refer to the theme of evil, works of Hawthorne, and Hawthorne himself. Such topical diversity helps explain the decrease in reference ties from biographical to theme texts and may even explain the increase in lexical ties used in the theme text condition.

The same sort of phenomenon is apparent in looking at the conjunctive ties used in the essays. In the biographical essays, the majority (67%) of the conjunctive ties consist of such expressions as during the first part of Hawthorne's life, as a child, as he grew up, after college, in 1839, after his resignation, etc. occurring in sentence initial position. They represent a class of conjunctive clauses and phrases adapted from Halliday and Hasan's conjunctive categorization scheme. With the theme of evil topic the bulk (75%) of the conjunctive ties consisted of such expressions as in his short stories, in "Rappacini's Daughter," in his novels, in "The Scarlet Letter" etc., also occurring at the beginning of sentences. The conjunctive ties used for the biographical topic are predominantly ones that orient the reader with respect to a time in Hawthorne's life. The conjunctive ties used for the theme topic are predominantly ones that orient the reader with

respect to Hawthorne's work. Here, there are decided differences in conjunctive usage related to the difference in text topic. But this is no surprise. We intuitively understand that the use of different types of conjunctive ties is a product (not a determinant) of a coherent rendering of a given topic. It is the topic difference that determines the difference in the type and proportions of conjunctive ties used for each essay.

The point is that gross cohesive patterning in text manifests itself in a strong main effect for text topic. One can only conclude from this that topic elicits an effect on the kinds of cohesive ties writers use; one cannot conclude that there is a topic effect related to coherence.

This negative conclusion bears directly on our question: Is a statistical accounting of cohesive ties a legitimate means of evaluating textual coherence? Because there was no significant variation due to familiarity in cohesive patterning amongst essays written on the same topic, whether the biographical or theme topic, there can therefore be no relationship between cohesive patterning and the coherence ratings of the essays within each topic.

Results of a Comparison of Familiarity Condition, Coherence Ranking, and Cohesive Patterning Features

The conclusion about a nonsignificant familiarity effect on cohesion may be premature, for a Familiarity x Topic interaction seems to surface when examining particular aspects of cohesive patterning. For this reason, in order to answer the question about the relationship between familiarity

and coherence we will continue to look at cohesive patterning. Furthermore, although the familiarity condition did not generate differences relative to cohesion, we do not know whether or not familiarity conditions might be tied to the coherence rankings of the essays written to biographical or theme topics. As we will see, the answer appears to be positive.

In order to investigate the relation of the variables of Familiarity, Coherence Ranking, and Cohesive Patterning, several aspects of topic-specific cohesive patterning were examined and compared with the coherence rankings, with each essay tagged with respect to its familiarity condition. The following features of topic-specific patterning were examined: First, we established a proportion based on the ratio of pronouns (P) and lexical (L) ties used in referring to Hawthorne to total (T) cohesive ties for the essay. This ratio will be abbreviated (P + L)/T. These features were chosen for examination because they correspond to the types of features expected in a biographical essay (see Morgan and Sellner's 'Churchill' example above). Also, they incorporate the significant contribution of the reference and lexical cohesion categories to the topic main effect. These proportions are given in Table 3. Table 4 presents the group means and standard deviations of the transformed proportion values. An ANOVA was conducted on this data and reflected the same trend established for the MANOVA. That is, there was a significant text-topic effect, $F(1,20) = 133.16$, $p < .01$, accounting for 86% of the variance. There was no interaction effect,

$p = .35$, nor was there a main effect due to familiarity, $F(1,20) = .03$, $p = .87$.

Second, proportional values were calculated for temporal conjunctives (TC) represented as a proportion of the total (T) conjunctive ties used for each essay. This ratio will be abbreviated TC/T. The temporal conjunctives were looked at because they were the most prominent type of connectives used in the essays, and, as pointed out above, they also play a role in determining the topic effect revealed by the MANOVA. These individual proportion values are also given in Table 3. An ANOVA was conducted on the transformed values of these proportion scores and revealed a Familiarity x Topic interaction, $F(1,20) = 21.37$, $p < .01$, accounting for 33% of the variance. The group and text main effects were also found to be significant, with Familiarity, $F(1,20) = 12.34$, $p < .01$, accounting for 19% of the variance and Topic, $F(1,20) = 10.84$, $p < .01$, accounting for 17% of the variance. Table 4 includes the group means and standard deviations used for this ANOVA.

To assess whether coherence corresponded with the cohesive patterning features examined, the rankings for the two essay topics were obtained as described in the rankings section. Tables 5 and 6 summarize the ranked data for each topic. The two rankings correspond to separate rankings of the essays written on the biographical topic and the essays written on the theme topic. The rankings are compared with the non-transformed proportion values of the two cohesive patterning features described above.

These values and the familiarity condition are given for both topic conditions. For two texts in each topic condition, the raters of text coherence did not agree on their ranking. The data on these texts are included in parentheses at the bottom of Table 5. These texts were not considered in the following discussion.

Discussion of the Comparison of Familiarity Condition, Coherence Rankings, and Cohesive Patterning Features

It must be realized that the reason for working with cohesive patterning features as opposed to the general cohesive categories is twofold. First, as mentioned previously, these features informally test out points in Morgan and Sellner's argument. Second, the proportion values for the general cohesive categories are gross values. They say little more than that there are patterning differences between topics. Using the patterning features we have chosen gets at the nature of how patterning actually manifests itself within topic.

With respect to the proportion figures in Table 3, and the observed interaction in particular, several points need to be made. The lower proportion values for temporal conjunctives used in the Familiar x Theme Topic cell stem from the familiar writers' use of additive conjunctives while working on the topic of the theme of evil in Hawthorne's work. This topic, in contrast to the biographical topic, can be described as more abstract and as developed through examples from Hawthorne's short stories and novels. The dominant additive conjunctives used are also and and.

The use of these conjunctives by the familiar group writing on the abstract topic does not seem a significant finding at this point. This is because the actual average number of uses of an additive conjunct per text in the Familiar x Theme Topic group is only two. A much larger sample size is needed to confirm the notion that a familiarity factor plays a part in determining specific uses of conjunctive tie-types.

However, a note of caution--the Familiarity x Theme Topic interaction should not be regarded as irrelevant. If one is searching for indications of the effect of prior knowledge on expression this is a good place to start. Conjunctives may show up as clear signals of logical relationship between ideas--relationships better understood by the familiar group. Also, after the writing tasks were completed, the experimenters became aware of the fact that the outlines--provided every student--possibly nullified the effects of the familiarizing filmstrip on the life and work of Hawthorne. Prior knowledge is too important a concept to ignore. In the present study, the experimenters concluded that an artifact of the design (using the outlines) lessened the possibility that any clear familiar/prior knowledge effect might surface.

The (P + L)/T Proportions. As for the comparison of the coherence rankings with the proportion values, as detailed in Table 5, several points can be made. The (P + L)/T proportions in the biography topic condition show a moderate negative correlation with coherence ranking ($r = -.47$). The theoretical position on the function of cohesive ties explored in this paper does not claim a relationship between high coherence level and low

number of cohesive ties. It may be the case that the writers in this study were using some other means of rendering a coherent text that superceded the use of cohesive ties. However, such a variable is not identified in this study. To the extent that an r of -.47 is considered strong, an argument can be made against the general claim that a large number of cohesive ties predicts a text's coherence. Also, with respect to the biography texts, there is no pattern of group designation as one reads down the rankings--no familiarity effect appears latent in these figures. The rank order correlation between group designations and coherence rankings was low (r = .12).

However, for the proportions presented in the theme text condition there is a strong negative correlation of group designation with coherence ranking (r = -.72). The five top ranked essays all belong to the familiar group (out of six possible). Here it is obvious there is a strong familiarity effect correlating with relative coherence. This effect does not show up in the MANOVA since those analyses are based on counts of cohesive ties--counts deemed non-significant. This non-significance is apparent in the rank order correlation of the (P + L)/T proportions with coherence ranking in the theme text condition (r = -.14).

The suggested familiarity effect indicates that perhaps information in the outline for the theme topic was more difficult to extract from the biographical outline. In sum, the familiarity effect suggests that there is an interaction of Familiarity x Text topic when looking at coherence

but not when looking at the (P + L)/T cohesion proportions. The main point of suggesting these effects and interactions is to show that crucial notions of familiarity, topic, and coherence seem in no way related to the specifically linguistic aspect of texts detailing the use of lexical and reference ties to refer to Hawthorne.

The TC/T Proportions. The TC/T Proportions also suggest effects not found in the MANOVA for the cohesive types. In both topic conditions, the top ranked essays all reveal a certain variation in the use of conjunctive tie-types. In other words, out of the total set of conjunctive ties used in the top-ranked essays in both text conditions, there were at least two conjunctive tie-types used, as opposed to the strict use of temporal conjunctives in the lower ranked essays (proportion values = 1.000). Unfortunately, in the TC/T proportions for the biographical essays this variation, although observable in terms of the proportion values, is not obvious in the essays themselves. In other words to get proportion values of .933, .833, .778, .917, and .857, usually only one non-temporal conjunctive need be used. This kind of variation can hardly be valuable in comprehension research, although it may prove a statistically reliable variation if examined over a large number of essays.

In the Familiar x Theme Topic cell, the situation in terms of the proportion values is only slightly different. As noted informally above, the TC/T proportions for the biography and theme texts, as opposed to the (P + L)/T proportions, strongly correlate with coherence rankings (r of

TC/T with the coherence rankings of the biography texts is -.67, r of TC/T with the coherence rankings of the theme texts is -.68). Also, it was noted previously that the Familiar Group x Theme Text cell used a certain proportion of additive tie-types. This finding was disclaimed because of the small n, the small number of actual additive ties used per essay (approximately 2), and the typical additive conjunctives used (and, also). However, it was suggested that if one wanted to establish a familiarity effect, or a familiarity effect, or a Familiarity x Topic interaction measurable by a cohesion variable, then this was perhaps a place to start. The observed cohesive patterning suggests that with the more difficult text topic, such an interaction effect may exist. It suggests that the topic might demand use of a variety of conjunctives in order to establish complex or varied relationships within the specific topic. This suggestion is cursorily investigated below by examining the textual context of the additive conjunctives used in the five top-ranked theme essays. Our findings indicate that even at this level of text analysis the TC/T proportions, representing the cohesion factor, do not seem related to a Familiarity x Topic interaction. Thus the coherence factor, at the base of the Familiarity x Topic interaction cannot be explained by the cohesion variable.

Observations on the use of additive conjunctives in text. Appendix D presents the textual context of additive conjunctives used in a particular section of the essays written by the Familiar x Theme Topic group, and provides all the alternative expressions used in relating the identical

information conjoined by the additive conjunctives. The relevant sections of the transcript and the outline abstracted from the transcript for the particular section are presented in Appendix C.

It appears that the use of an additive conjunctive at this point in a student's essay is as much a product of the structural characteristics of the outline, as it is of a fuller understanding of the material. In all the texts, except for the first and seventh ranked essays, the relationship between the Puritan zealots and the hedonistic May-pole revellers is correctly stated. In other words, in these essays the notion that whatever-the-zealots-did-the-revellers-did-too is understood.

What appears as not well understood by the writers of the lower ranked essays is the nature of the relation between potential virtues of the zealots and revellers and the groups' corruption. The sixth and seventh ranked essays equate corruption, degeneracy, and fanaticism with potential virtues. The eighth and ninth ranked essays use extremely awkward and inappropriate expressions (equalizes and specifically) in establishing the relationship between the zealots and the revellers. And the tenth ranked essay confuses the meaning of hedonistic, using notions of happiness and joy as the principle vehicles for the corruption of the revellers. Minimally, it may be suggested that any confusion within the outline was resolved for the familiar group because they were somewhat familiar with the concepts communicated.

More importantly, it is not the case that the use of an additive conjunctive contributed in any direct way to the expression of the

concepts presented in the outline. Except for the first ranked essay which generalizes across zealots and revellers, and the seventh ranked essay which misstates the relation of zealots to revellers, each essay uses expressions which correctly state what is essentially a structural relation between the Puritan zealots and the May-pole revellers in the outline. These expressions include the additive conjunctives also, and, and again; the comparative reference items similar and same; the conjoining expressions becoming as, and equalizes; a use of the colon; and the construction of the generalized subject the two main groups. All of these expressions achieve the same effect and in no way distinguish, either in terms of style or level of coherence, sections of essays which vary radically in the degree to which they correctly interpret other information in the outline.

A similar argument can be made for other contexts of use of additive conjunctives within the top ranked five essays of the theme text condition. Nor is it more constructive to investigate the nature of the variation in the use of conjunctive items in the biography text condition. In other words, at this level of cohesive text analysis, the cohesion index is an inadequate predictor of coherence.

On coherence: An example of a coherent rendering of the theme of evil topic. At this point it must be asked, is there anything positive to say about the relationship of the use of cohesive items in an essay and that essay's general coherence level? The answer seems to be: Only insofar

as the ties are used unambiguously. The point is, to the extent that the writer constructs a coherent rendering of the content of the outlines will the tie elements unambiguously signal relationship and reinforce conceptual expectations of the text.

For example, making a strictly subjective evaluation of a 'best' text in Appendix 4, consider the fifth ranked essay's statement in a larger textual context;

> In the story "May-Pole of Merry Mount," the Puritan zealots carried to extremes their potential virtues. In this way the zealots became corrupt, degenerate, and fanatical. The hedonistic May-pole revellers carried their potential virtues to an extreme, becoming as the Puritan zealots. Here Hawthorne points out man's struggle against potential evil.

From a more global processing perspective, what makes this paragraph "work" is the fact that this writer fully understands the terms, and the relationship between the terms, in the argument presented by the outline. This writer understands that the most superordinate concept in the outline is the theme of man's potential for evil as it is expressed in the works of Nathaniel Hawthorne. The writer understands that within this concept the theme of man's potential for evil is primary throughout the outline, while reference to various works of Hawthorne is secondary to the extent that they are a means of developing the general theme of man's potential for evil. The writer frames his paragraph with this understanding, choosing to use the cohesive expressions <u>in the story</u> and <u>here</u> to focus on the "May-Pole

of Merry Mount" and to use the lexically cohesive expression man's struggle against potential evil as an instance of the articulation of the major theme.

From a more local processing perspective, the writer fully understands the terms, and the relationship between the terms, in the subordinate arguments presented in the paragraph on the "May-Pole of Merry Mount." The writer understands the subordinate concept that the Puritan zealots carried their potential virtues to such extremes that as a group they became corrupt. The writer understands that this subordinate argument generalizes to the case of the May-pole revellers. The writer signals this argument generalization by using the non-cohesive phrase becoming as. And he understands that the parallel drawn between the Puritan zealots and the May-pole revellers constitutes the narrative means by which Hawthorne made the "May-Pole of Merry Mount" a comment on man's potential for evil.

The writer's understanding of the outline, reflected in the coherence rankings and in his actual writing is certainly not to be conceived as a product of the cohesive items. The main cohesive items used are more readily understood in terms of their epiphenomenal role rather than in any coherence-producing role. For example, as discussed above, the use of what was labelled as conjunctive phrases orienting the reader with respect to Hawthorne's work (in the story, here, etc.) follows from the topic of the essay--the theme of evil in the works of Hawthorne. Similarly, the unambiguous use of the lexically cohesive elements constituting the phrase man's struggle against potential evil again follows directly from an

understanding of the topic for the essay--the theme of evil in the works of Hawthorne. In summary, the answer to the question of the relationship of the use of unambiguous cohesive elements to text coherence is fairly straight-forward in this study. Cohesion follows from a coherent rendering of a text by the writer.

But if cohesion is not to be regarded as an index of coherence, is there any objective measurement that might predict the superior coherence for a reader of the sample just discussed? This question cannot be answered based on the information presented here. But it can be said, based on the above discussion of the "best" text, that several conditions must hold for relative coherence: (a) an overall structure permeating the text as a function of the argument of the text; and (b) the signalling of relationships between terms of the argument (there are a variety of means, cohesive ties being only one). It appears that comprehensibility in text might best be approached from an argumentation perspective. It is essential, though, that the argument not be misinterpreted as structure. In other words, the text need not be architecturally bound from beginning to end, but the status of the argument must be comprehensible at any point in a text.

Conclusions and Research Implications

With the analyses that have been done with counts of types of cohesive ties, little positive has been stated about the causal relation of Halliday and Hasan's cohesion concept to textual coherence. What was found is that

cohesive ties are pervasive in text and are patterned across topics. But ties are pervasive almost by definition since reference, conjunction, and lexical cohesion include a large proportion of any text (over the total 24 essays there is an average of 92 ties per essay with each tie consisting of a minimum of 2 words which means a minimum of 184 words function as tie elements per essay, with each essay averaging 310 words). Such pervasiveness severely diminishes the usefulness of the cohesion concept as an index of coherence at a global or local level. This statement does not contradict the significant results for text conditions found in the MANOVA. Rather, the point is that this finding is too general to be of any use. The topic effect is real, but appears as the product of a gross measurement of cohesion and represents intuitive notions about differences between texts written on different topics.

With respect to our general question, this study argues against using cohesion analysis as an index or predictor of a text's coherence. There appears to be no causal relationship between proportional measures of cohesive ties within topic and coherence rankings within topic. In other words, the present study indicates that a cohesion index is causally unrelated to a text's coherence. Interesting results seem to be those primarily concerned with topic, familiarity, and coherence ranking.

Overall, based on the findings of this study, it can be said that statements about the function of cohesive elements in coherent texts (a) should not be based on a statistical accounting of ties using group means as a basis for comparing groups; (b) should not be based on a statistical

accounting of ties within single texts--i.e , the cohesion concept should not be used to characterize individual texts, as if the totality of cohesive ties represents a cumulative macro-effect on text coherence; and (c) should be based on a description of single cohesive instances within a text viewed as consequences of a coherent rendering of the argument of a text.

References

Halliday, M. A. K., & Hasan, R. Cohesion in English. London: Longman, 1976.

Kirk, R. E. Experimental design: Procedures for the behavioral sciences. Belmont, Cal.: Brooks/Cole, 1968.

Levy, D. M. Communicative goals and strategies: Between discourse and syntax. In T. Givon (Ed.), Syntax and semantics (Vol. 12): Discourse and syntax. New York: Academic Press, 1979.

Meyer, B. J. F. The organization of prose and its effects on memory. Amsterdam: North-Holland, 1975.

Morgan, J. Towards a rational model of discourse comprehension. TINLAP-2, 1978.

Morgan, J., & Sellner, M. Discourse and linguistic theory. In R. J. Spiro, B. C. Bruce, W. F. Brewer (Eds.), Theoretical issues in reading comprehension. New York: Erlbaum, 1980.

Tierney, R. J., & Mosenthal, J. H. Discourse comprehension and production: Analyzing text structure and cohesion (Tech. Rep. No. 152). Urbana: University of Illinois, Center for the Study of Reading, January 1980. (ERIC Document Reproduction Service No. ED 179 945)

Appendix A

Outline for the Topic of the Theme of Evil in the Works of Hawthorne

TOPIC	DETAILS
evil	theme
	NH's work
NH	ambivalent
	evil
ambivalence	Puritans' premise
	man capacity for evil
	counter to Transcendentalist Movement
	man good
	theme
	NH's work
"May-Pole of Merry Mount"	evil potential
	man
Puritan zealots	potential virtues
	carried to extremes
	corrupt
	degenerate
	fanatical
hedonistic May-pole revellers	potential virtues
	carried to extremes
	corrupt
	degenerate
	fanatical
"Rappaccini's Daughter"	deceptions
	evil
garden	
fantastic	location
	story
	beautiful
	poisonous
	tragic
	symbolic parallel
	heroine

TOPIC	DETAILS
"Young Goodman Brown"	ambivalence nature beautiful to NH evil beneath surface puritanical awareness
The Scarlet Letter	story effect of sin 3 members of New England community Hester Prynne adultery Reverend Dimmesdale hypocrisy Chillingsworth violation of human heart "unpardonable sin"
The Marble Faun	resolves ambivalence ambitious ambiguous NH's last novel reenactment Fall of Man fortunate man sin knowledge of presence of evil become aware of morality

Outline for a Biographical Sketch of Hawthorne

TOPIC	DETAILS
NH	born July, 1804
	Salem, Mass.
	recluse
	until marriage
recluse	
child	solitude of woods and fields
college	Bowdoin College, Maine
after college	Salem
	with widowed mother
	2 spinster sisters
until (NH's) mid-thirties	desire for solitude
1839	job
	Boston Custom House
	supervisor
	goods from merchant ships
	enjoyed work
	later
	detest contact with "world"
April, 1841	Brook Farm
	Transcendentalists' utopian community
	change
	privacy
	didn't work
	left 1842
marriage	Sophia Peabody
	respected family Salem
	30 years old
	desire for solitude
	seem to do away with

TOPIC	DETAILS
3 years	Salem from Concord
	happy
	productive
	less and less a recluse
productive and social period	end
	financial need
job	custom house
	resign 1849
	wife's frugality
write	burst of creativity
	The Scarlet Letter
	success
	demands
	seclusion
	Lenox, Mass.
	temporary
Concord	write
7 years	post overseas
	from President Franklin Pierce
	college friend
	with family
May, 1864	death
	Concord
	at peace with himself and world

Appendix B

Sample Text with Cohesion Analysis

STUDENT ESSAY

The Life of Nathaniel Hawthorne

Nathaniel Hawthorne, the great American writer, was born in July, 1804. He lived in Salem, Massachusetts. During the first part of Hawthorne's life and until the time of his marriage, he tended to be a withdrawn type of person.

As a child, Hawthorne found solitude in the woods and fields and often went there to enjoy that solitude. As he grew up, he attended college at Bowdoin College in Maine. After college, Hawthorne returned to Salem and lived with his widowed mother and two spinster sisters. The solitude that Hawthorne enjoyed as a child was still one of his great desires even until his mid-thirties.

COHESIVE ANALYSIS OF ESSAY

Tie #	Sentence	Paragraph	Cohesion Category	Presupposing Item	Presupposed Item
1	1	1	Lexical	Nathaniel Hawthorne	Nathaniel Hawthorne
2	2	1	Reference	He	Nathaniel Hawthorne
3	3	1	Lexical	Hawthorne's	Nathaniel Hawthorne
4	3	1	Lexical	Life	Life
5	3	1	Conjunctive	During (the first part of Hawthorne's life) and until (the time of his marriage)	Sentence 2
6	4	2	Lexical	Child	Person
7	4	2	Conjunctive	As a child	Sentence 3
8	4	2	Lexical	Hawthorne	Hawthorne's

Tie #	Sentence	Paragraph	Cohesion Category	Presupposing Item	Presupposed Item
9	4	2	Lexical	Solitude x 2	Withdrawn
10	5	2	Reference	He x 2	Hawthorne
11	5	2	Conjunctive	As he grew up	Sentence 4
12	5	2	Lexical	Maine	Massachusetts
13	6	2	Conjunctive	After college	Sentence 5
14	6	2	Lexical	College	College
15	6	2	Lexical	Hawthorne	Hawthorne
16	6	2	Lexical	Salem	Salem
17	6	2	Lexical	Lived	Lived
18	7	2	Reference	The (solitude)	Solitude x 2
19	7	2	Lexical	Hawthorne	Hawthorne
20	7	2	Lexical	Enjoyed	Enjoy
21	7	2	Lexical	Child	Child

Appendix C

Outline and Original Text for Text Passage Generating Occasional Usage of Additive Conjuncts

OUTLINE

Topic	Detail
Puritan zealots	potential virtues
	carried to extremes
	corrupt
	degenerate
	fanatical
hedonistic May-pole revellers	potential virtues
	carried to extremes
	corrupt
	degenerate
	fanatical

ORIGINAL TEXT

The overbearing Puritan zealots as well as the hedonistic May-pole revellers carry their potential virtues to such extremes that both groups are depicted as corrupt, degenerate, and fanatical.

Appendix D

TEXTUAL CONTEXT OF ADDITIVE CONJUNCTS

Rank	Group	Text (sentences with additive conjuncts are starred)
1	Familiar	In "May-Pole of Merry Mount," Nathaniel Hawthorne wrote about man's potential for evil and showed how Puritans believed that man's potential virtues were corrupt, degenerate, and fanatical.
2	Familiar	*He shows also, these hedonistic May-pole revellers, which shared similar potential virtues,
3	Familiar	*Their belief in their virtues also led them to become corrupt, degenerate, and fanatical,
4	Familiar	*He also included the pleasure seeking May-pole revellers to be the same: their potential virtues were carried to extremes. Again, man was corrupt, degenerate, and fanatical.
5	Familiar	The hedonistic May-pole revellers carried their potential virtues to an extreme, becoming as the Puritan zealots.
6	Unfamiliar	The two main groups in the story, the Puritan zealots and the pleasure seeking May-pole revellers, carry their potential virtues, corruption, degeneration, and fanaticism, to the extremes.
7	Unfamiliar	He showed how Puritan zealots carried corrupt, degenerate, and fanatical acts, which were their potential virtues, to extremes by using hedonistic May-pole revellers.
8	Unfamiliar	Hawthorne equalizes the zealots with the hedonistic May-pole revellers whose virtues took them to the same fates of corruption degeneration and fanaticism.
9	Familiar	The Puritan zealots carried potential virtues to the extremes. They show corruption, degeneracy, and fanaticism. Specifically in the novel the hedonistic May-pole revellers had the same characteristics.
10	Unfamiliar	*On the other side of the zealots are the hedonists who through their happiness and joy also carry the virtues to extremes, again to corruption, to degeneracy, and to almost fanatical states.

Table 1

Proportion Data for Cohesive Analyses of Texts

Student Number	Biography Text			Theme Text		
	Reference	Conjunctive	Lexical	Reference	Conjunctive	Lexical
			Familiar			
1	.330	.080	.591	.200	.057	.743
2	.222	.049	.728	.082	.082	.837
3	.383	.083	.533	.109	.063	.828
4	.135	.146	.802	.190	.091	.719
5	.200	.150	.650	.122	.073	.805
6	.250	.138	.612	.118	.067	.815
M	.253	.108	.653	.137	.072	.791
SD	.090	.042	.098	.047	.013	.048
			Unfamiliar			
7	.284	.095	.600	.110	.044	.846
8	.267	.143	.590	.108	.032	.860
9	.190	.076	.734	.081	.070	.849
10	.250	.075	.675	.125	.014	.861
11	.323	.075	.602	.133	.060	.807
12	.080	.120	.800	.056	.045	.899
M	.232	.097	.667	.102	.044	.854
SD	.086	.028	.086	.029	.020	.030

Table 2

Group Means and Standard Deviations of the Transformed Scores

Group	Text Condition						Total Group		
	Biography Text			Theme of Evil Text					
	Reference	Conjunctive	Lexical	Reference	Conjunctive	Lexical	Reference	Conjunctive	Lexical
Familiar									
M	1.044	.657	1.888	.750	.536	2.196	.897	.597	2.042
SD	.209	.143	.212	.136	.055	.117	.228	.121	.249
Unfamiliar									
M	.990	.636	1.917	.645	.413	2.360	.818	.524	2.138
SD	.228	.094	.188	.101	.106	.085	.247	.151	.270
Total Text									
M	1.017	.647	1.902	.697	.475	2.278	-	-	-
SD	.211	.116	.191	.126	.103	.130	-	-	-
Total									
M	-	-	-	-	-	-	.857	.561	2.090
SD	-	-	-	-	-	-	.236	.139	.250

Table 3

Proportions for Types of Cohesive Patterning

Student Number	Text Condition			
	Biography Text		Theme Text	
	(P + L)/T[a]	TC/T[b]	(P + L)/T	TC/T
	Familiar			
1	.341	1.000	.152	.667
2	.296	1.000	.133	.500
3	.350	1.000	.063	.750
4	.250	.857	.132	.727
5	.270	.933	.122	.833
6	.293	.937	.067	.500
	Unfamiliar			
7	.263	.778	.121	1.000
8	.286	1.000	.151	1.000
9	.304	.833	.081	.833
10	.300	1.000	.111	1.000
11	.312	1.000	.145	1.000
12	.220	.917	.101	1.000

[a]Proportion of (Pronouns and Lexical Ties Referring to Hawthorne)/(Total Number of Ties for Essay) = P + L/T.

[b]Proportion of (Temporal Conjunctives)/(Total Conjunctive Ties for Essay) = TC/T.

Table 4

Group Means and Standard Deviations of Transformed Values for $(P + L)/T^a$ and TL/T^b Proportions

Group	Text Condition				Total Group	
	Biography Text		Theme Text			
	(P + L)/T	TC/T	(P + L)/T	TC/T	(P + C)/T	TC/T
			Familiar			
M	1.158	2.636	.668	1.916	.913	2.276
SD	.085	.138	.122	.293	.275	.435
			Unfamiliar			
M	1.114	2.535	.699	2.056	.907	2.596
SD	.077	.249	.087	.178	.232	.216
			Total Text			
M	1.136	2.586	683	2.286	.910	2.436
SD	.084	.199	.102	.450	.249	.373

[a](Pronouns + Lexical)/(Total) = (P + L)/T

[b](Temporal Conjunctives)/Total Conjunctives) = TC/T

Table 5

Coherence Rankings and Aspects of the Cohesive Analyses for the Biography Texts

Rank	Familiarity Condition*	Proportions	
		$(P + L)/T^a$	TC/T^b
1	F	.293	.937
2	U	.304	.833
3	U	.263	.778
4	F	.296	1.000
5	U	.220	.917
6	F	.250	.857
7	U	.286	1.000
8	F	.341	1.000
9	U	.312	1.000
10	F	.350	1.000
**	(F)	(.270)	(.933)
**	(U)	(.300)	(1.000)

[a]Lexical and reference ties used to refer to Hawthorne expressed as proportion of total cohesive ties for essays.

[b]Temporal conjunctives expressed as a proportion of total conjunctive ties used for essay.

*Group designation: F = familiar, U = unfamiliar.

**Text not used in comparisons between coherence ranking, cohesion proportions, and familiarity conditions.

Table 6

Coherence Rankings and Aspects of the Cohesive Analyses for the Theme Texts

Rank	Familiarity* Condition	Proportion	
		$(P + L)/T^a$	TC/T^b
1	F	.063	.750
2	F	.152	.667
3	F	.133	.500
4	F	.132	.727
5	F	.067	.500
6	U	.101	1.000
7	U	.151	1.000
8	U	.111	1.000
9	F	.122	.833
10	U	.145	1.000
**	(U)	(.121)	(1.000)
**	(U)	(.081)	(.833)

[a]Lexical and reference ties used to refer to Hawthorne expressed as proportion of total cohesive ties for essay.

[b]Temporal conjunctives expressed as a proportion of total conjunctive ties used for essay.

*Group designation: F = familiar, U = unfamiliar

**Text not used in comparisons between coherence rankings, cohesion proportions.